

Geospatial Analysis in R

Anjile An [jia4001@med.cornell.edu]

Why map data?

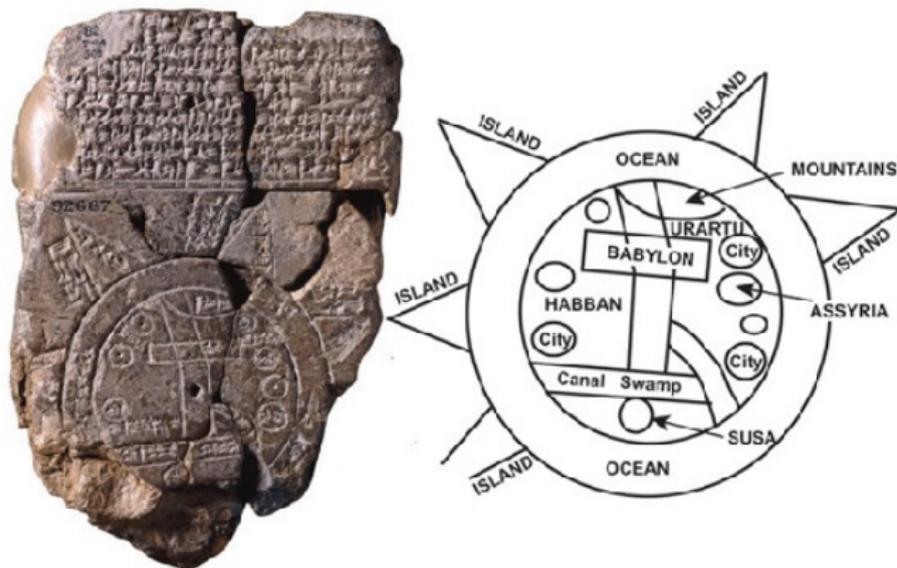


Figure 1: Imago Mundi: Babylonia 500 BCE

Why map data?

- ▶ Maps simplify complicated information
- ▶ Maps are functional tools - show where resources are, natural features
- ▶ Maps give stories context
- ▶ Maps are visual!
- ▶ Etc etc etc

Data + Maps

What insights can maps show us?



Figure 2: Broad Street Cholera Map, 1854

Data + Maps

Maps with data are particularly powerful:

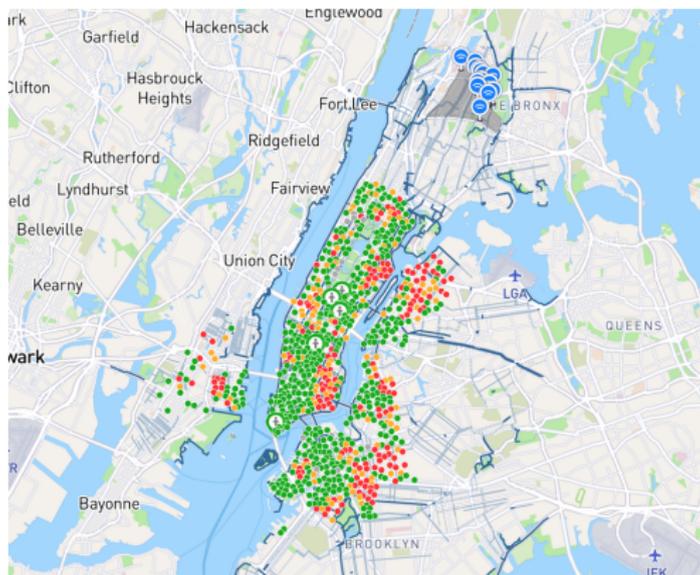


Figure 3: NYC Citibike Docks, 2019

What is GIS?

A Geographic Information System (GIS) links locational (spatial) and database (tabular) information and enables a person to visualize patterns, relationships, and trends. This process gives an entirely new perspective to data analysis that cannot be seen in a table or list format. The five components of a GIS are listed below.

HARDWARE

The hardware is the computer and peripherals on which the GIS operates. Today, this could be a centralized computer server running the UNIX or Windows NT operating systems, a desktop PC, or an Apple Macintosh. The computer may operate in isolation or in a networked configuration.

- Computers
- Networks
- Peripheral Devices
 - Printers
 - Plotters
 - Digitizers



SOFTWARE

GIS software provides the functions and tools users need to store, analyze, and display geographical information. The key software components are:

- GIS Software
- Database Software
- OS Software
- Network Software



DATA

One of the most important components of GIS is the data. It is absolutely essential that data be accurate. The following are different data types:

- Vector Data
- Raster Data
- Image Data
- Attribute Data



GIS

PEOPLE

GIS technology is clearly of limited value without people to manage the system and to develop plans for applying it. Users of GIS range from highly qualified technical specialists to planners, foresters, and market analysts who use GIS to help with their everyday work.

- Administrators
- Managers
- GIS Technicians
- Application Experts
- End Users
- Consumers



METHODS

Methods are well designed plans and application-specific business rules describing how technology is applied. This includes the following:

- Guidelines
- Specifications
- Standards
- Procedures



GIS Specific Software

R can do a lot of geospatial analysis, but there are lots of things it can't do. Here are some GIS software I've used:

- ▶ ArcGIS (paid - "gold standard")
- ▶ QGIS (open source)
- ▶ Google Earth Pro
- ▶ GeoDa

Types of Geospatial Data

There are two main types, raster and vector.

- ▶ Vector data has dominated the social sciences since many features used tend to have discrete boundaries
- ▶ Raster is utilized most in environmental sciences since remote sensing data comes in as images which are more continuous in nature.

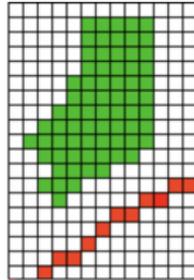
Types of Geospatial Data



Real World



Vector

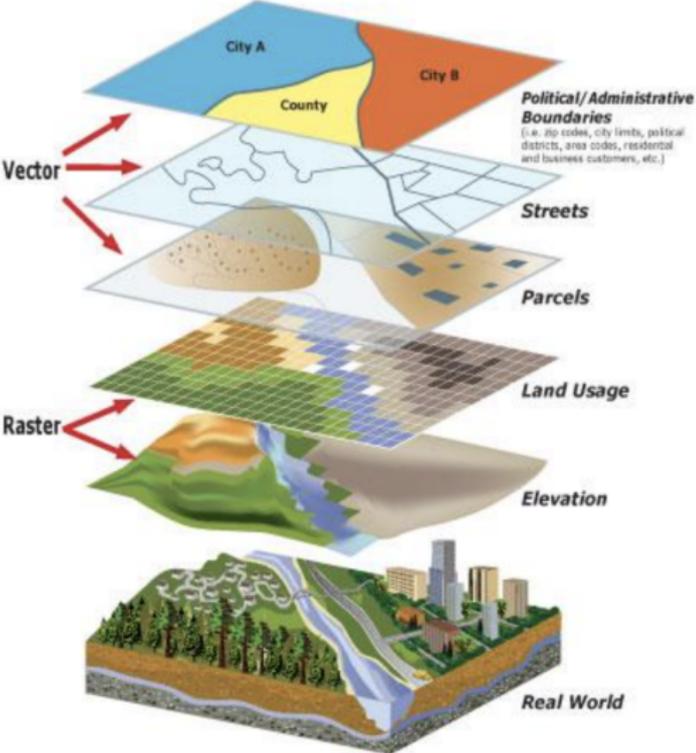


Raster

Types of Geospatial Data

Data format	Used with...	Used in package...	Used for...	Comment
<code>sf</code>	vector	<code>sf</code> , others	visualizing, manipulating, querying	This is likely to become the new spatial standard in R. Will also read from spatially enabled databases such as postgresSQL.
<code>raster</code>	raster	<code>raster</code> , others	visualizing, manipulating, spatial statistics	This is the most versatile raster format
<code>SpatialPoints*</code> <code>SpatialPolygons*</code> <code>SpatialLines*</code> <code>SpatialGrid*</code>	vector and raster	<code>sp</code> , <code>spdep</code>	Visualizing, spatial statistics	May be superseded by <code>sf</code> in the future
<code>ppp</code> <code>owin</code>	vector	<code>spatstat</code>	Point pattern analysis/statistics	NA
<code>im</code>	raster	<code>spatstat</code>	Point pattern analysis/statistics	NA

Layered Model



Types of Maps

- ▶ Physical maps show landscape features like rivers and land masses
- ▶ Topographic maps show the contours of elevation lines
- ▶ Climate maps can show atmospheric conditions
- ▶ Thematic maps show special topics like election results or even energy usage in areas.

Coordinate Reference Systems

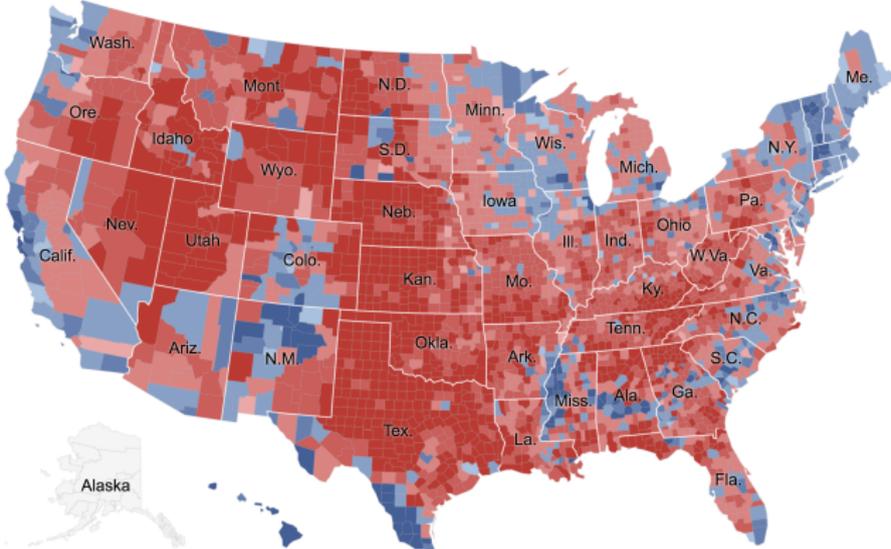
Why maps showing the same area look “distorted” - the coordinate reference system is crucial since it ties the vector and raster data types to a location on the earth (or other bodies).



Reference Map



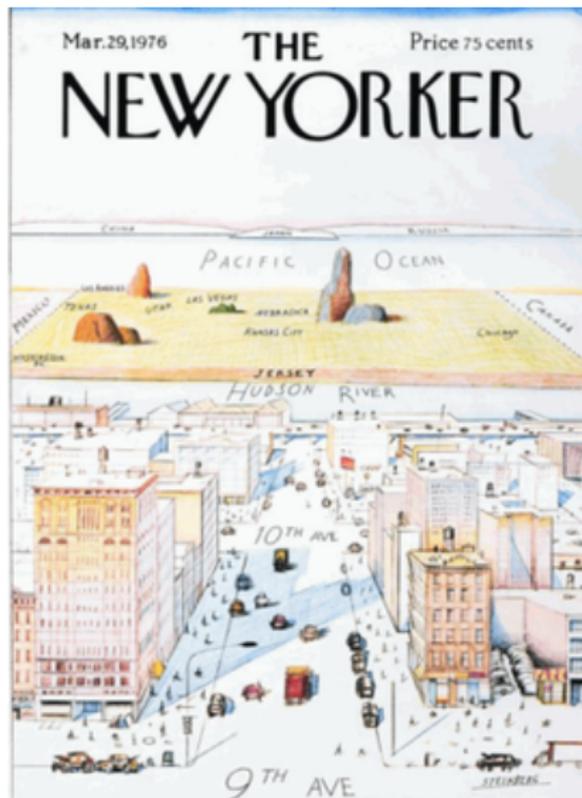
Chloropleth Map



3D Map



Mental Map



Interactive Map

The screenshot shows an interactive map interface titled "THE REFUGEE PROJECT". At the top, there are two tabs: "ORIGIN COUNTRY" (highlighted in orange) and "ASYLUM COUNTRY" (highlighted in teal). A back arrow is located between them. The main area features a dark world map with orange circles of varying sizes representing refugee populations in different regions.

On the left side, there is a data panel with the following content:

- REFUGEE POPULATION**
- 2018 ▾ World ▾
- 20,084,122**
- refugees from 156 countries in 152 asylum countries
- [Explore Data](#)
- A bar chart showing population distribution across countries.
- REFUGEE STORIES** (with UNHCR logo)
- 2018 PANAMA
- A photo of a woman and a man.
- Text: "Kigali sees economic sense in helping refugee entrepreneurs"
- LEARN WHY**
- 2018 CAMERSON

Navigation controls include a play button, a plus sign, and a minus sign.

Resources

This is just a very basic introduction! If you want to delve deeper:

- ▶ Geocomputation with R, Robin Lovelace, Jakub Nowosad, Jannes Muenchow (free, online)
- ▶ <https://www.r-spatial.org/>
- ▶ <https://www.r-graph-gallery.com/> (for inspiration)
- ▶ Google/stackoverflow

Geospatial Data Sources

- ▶ NYC OpenData
- ▶ ERSI
- ▶ TIGER GIS

Let's map!

Live coding session - download all the files in the Geospatial Analysis folder, and open up Geospatial_Analysis.R and install the mapping packages.

Let's map!

